CIGI

CHATHAM HOUSE
The Royal Institute of International Affairs

Global Commission on Internet Governance

ourinternet.org

PAPER SERIES: NO. 33 — MAY 2016

Market-driven Challenges to Open Internet Standards

Patrik Fältström

MARKET-DRIVEN CHALLENGES TO OPEN INTERNET STANDARDS

Patrik Fältström

CHATHAM HOUSE
The Royal Institute of International Affairs

Published by the Centre for International Governance Innovation and Chatham House.

This work was carried out with the aid of a grant from the International Development Research Centre (IDRC), Ottawa, Canada.

The views expressed herein do not necessarily represent those of IDRC or its Board of Governors.

CIGI

67 Erb Street West
Waterloo, Ontario N2L 6C2
Canada
tel +1 519 885 2444 fax +1 519 885 5450
www.cigionline.org

CHATHAM HOUSE
The Royal Institute of International Affairs

10 St James's Square
London, England SW1Y 4LE
United Kingdom
tel +44 (0)20 7957 5700 fax +44 (0)20 7957 5710
www.chathamhouse.org

TABLE OF CONTENTS

ABOUT THE GLOBAL COMMISSION ON INTERNET GOVERNANCE

The Global Commission on Internet Governance was established in January 2014 to articulate and advance a strategic vision for the future of Internet governance. The two-year project conducts and supports independent research on Internet-related dimensions of global public policy, culminating in an official commission report that will articulate concrete policy recommendations for the future of Internet governance. These recommendations will address concerns about the stability, interoperability, security and resilience of the Internet ecosystem.

Launched by two independent global think tanks, the Centre for International Governance Innovation (CIGI) and Chatham House, the Global Commission on Internet Governance will help educate the wider public on the most effective ways to promote Internet access, while simultaneously championing the principles of freedom of expression and the free flow of ideas over the Internet.

The Global Commission on Internet Governance will focus on four key themes:

- enhancing governance legitimacy — including regulatory approaches and standards;
- stimulating economic innovation and growth — including critical Internet resources, infrastructure and competition policy;
- ensuring human rights online — including establishing the principle of technological neutrality for human rights, privacy and free expression; and
- avoiding systemic risk — including establishing norms regarding state conduct, cybercrime cooperation and non-proliferation, confidence-building measures and disarmament issues.

The goal of the Global Commission on Internet Governance is two-fold. First, it will encourage globally inclusive public discussions on the future of Internet governance. Second, through its comprehensive policy-oriented report, and the subsequent promotion of this final report, the Global Commission on Internet Governance will communicate its findings with senior stakeholders at key Internet governance events.

www.ourinternet.org

ABOUT THE AUTHOR

Patrik Fältström is the head of Engineering, Research and Development at Netnod. He has previously served in leadership roles at IETF, the Internet Architecture Board and Internet Society (ISOC) (1998–2009); he was a distinguished engineer at Cisco Systems and has worked at Tele2 in Sweden, the Swedish Navy and Bunyip in Canada. Patrik was also appointed adviser to the Swedish IT minister (2003–2014), and has been the chair of the Internet Corporation for Assigned Names and Numbers Security and Stability Committee since 2011. Patrik has been working with Internet-related standardization and development since 1985 and is one of the founders of the ISOC Special Interest Group on Internet of Food.

Patrik holds an M.Sc. in mathematics from the University of Stockholm. He is a member of the Royal Swedish Academy of Engineering Sciences. In January 2011, Patrik received the Order of the Cross of Terra Mariana, V class, from the president of Estonia.

ACRONYMS

AIN	Advanced Intelligent Network
APIs	Application Programming Interfaces
CPE	customer premises equipment
DHCP	Dynamic Host Configuration Protocol
DNS	Domain Name System
HTTP	Hypertext Transfer Protocol
IEEE	Institute for Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IMAP	Internet Message Access Protocol
IoT	Internet of Things
IP	Internet Protocol
IRC	Internet Relay Chat
ISOC	Internet Society
ISPs	Internet Service Providers
ITU	International Telecommunication Union
ITU-T	ITU Telecommunication Standardization Sector
MIME	Multipurpose Internet Mail Extensions
NAT	network address translation
POP	Post Office Protocol
RFC	Request for Comments
SMTP	Simple Mail Transfer Protocol
TCP	Transmission Control Protocol
TLD	top-level domain names
W3C	World Wide Web Consortium
XMPP	eXtensible Messaging and Presence Protocol

EXECUTIVE SUMMARY

The success of the Internet as a dynamic foundation for building an enormous variety of interactive services depends on interoperability, open standards and the ability to innovate freely and provide services without permission, all of which arise from its edge-oriented architecture. Historically, these attributes have required private actors to leave the comfort zone of traditional telecommunications and embrace a regulatory environment that encourages change and innovation. But market forces emerging with new technologies are challenging these foundations, with potentially dire consequences for the continued openness of the Internet.

This paper establishes some of the basic Internet principles that have enabled innovation and interoperability, such as globally unique identifiers and open technical standards. It explains how market economy forces have shaped the evolution of Internet standards, including a resurgence of proprietary and anti-competitive approaches. The paper warns about specific market-driven trends that threaten to erode the Internet's openness, trends that will only accelerate in the context of the Internet of Things (IoT) and cloud computing. Finally, the paper makes recommendations for reversing the trend toward fragmentation through the use of open-standard protocols, the development of application programming interfaces (APIs) as if they were protocols, the use of open standards processes and the use of public procurement to encourage openness.

INTRODUCTION

The ideal of an "open" Internet is often portrayed as wishful thinking — something that would be nice to have in a perfect world, but is not always compatible with the need for revenue and the harsh reality of the market economy. This is a profoundly false impression, because the openness of the Internet and its mechanisms, from the development of technical standards to the operation of the global network, confers enormous practical economic benefits.

In practice, the open Internet has been fertile ground for the invention and development of remarkable new companies, capabilities and modes of human interaction. The openness principle continues to guide the Internet's evolution in technical, economic, political and social dimensions. Innovation in the open Internet is achieved by consensus through open collaboration among researchers, manufacturers, service providers and users. Innovation can start anywhere and propagate in any direction.

But that's the long-term view. In the short term, market forces can drive fragmentation and anticompetitive "silo" approaches to product and standards development that erode the foundations of the open Internet. These forces are not only short term with respect to commercial advantage but also short-sighted regarding sustainable innovation and economic growth. They can be countered by a clear understanding of the tangible benefits of the Internet's traditional open approach to standards development.

INTERNET FUNDAMENTALS

The Internet, like other communication networks from highways to telephones, consists of a web of connections, and anyone who is connected to the Internet can communicate with anyone else. The basic communication is via Internet Protocol (IP) packets, each of which carries a small amount of data; many of them together carry a document, a movie, a telephone call, a recipe or the latest photographs of someone's cat.

The point (or "node") at which each device is connected to the Internet is identified by an IP address, and because these addresses are globally unique, communication between two devices is, in its simplest form, a sequence of IP packets sent back and forth between them, in which each packet includes the IP addresses of both. The service we get with Internet access is therefore relatively simple: the best-effort delivery of IP packets from the source to whatever node has the address stated as the destination address in the packet.

The interpretation of the contents (payload) of each packet is up to the software in the nodes that send and receive the packets. This is why, when comparing traditional telecommunication and the Internet architecture, one says that the intelligence — the knowledge of what services exist — has moved from the network to the edge. Just by changing the software in two nodes that communicate (and without changing anything in the network), a new service can be launched. We call this "permissionless communication and innovation." Innovation and launch of new services has moved from being a business for the owner of the network to a business for whomever controls the devices connected to the Internet — the end user.

End-to-End Communication

The end-to-end communication between two Internet users is often referred to as "peer-to-peer," in part to distinguish it from "client-server" communication between an end user (the client) and a broadly available application service provided by someone else, such as Google, Netflix or Amazon (the server). From an Internet architecture point of view, these service companies also have nodes with unique IP addresses with which their customers (the end users) communicate. Both servers and their clients are connected to the Internet by companies that provide the infrastructure to move IP packets from one node to another. It is very important to distinguish the packet-level exchanges facilitated by these Internet service providers (ISPs) from the application-level services, such as "movie watching on demand," that are offered by content provider companies such as Netflix. In this example, both Netflix and its clients are separate customers of ISPs.

Global Uniqueness

The ability to exchange IP packets unambiguously between any two nodes on the Internet requires that all nodes have globally unique IP addresses. To make path selection easier, IP addresses are generally allocated according to network topology, so that two nodes that have addresses adjacent to each other are located within the same network. Because of this, IP addresses might change when a node is moving in the network. Domain names create longer-term stability in addressing, and make the addressing more human friendly. Each domain name is also globally unique, and the distributed database called the Domain Name System (DNS) maps domain names to IP addresses.

We Have One — and Only One — Set of Domain Names

Domain names are sequences of LDH-string[1] labels organized into a tree-structured hierarchy in which the manager of a domain at one level can allocate subdomains at the next level, moving away or "down" from the root of the tree. At the root, we guarantee the global uniqueness of all domain names by having one and only one manager of the top-level domain names (TLDs) (Request for Comments [RFC} 2826).[2] Because each domain name is unique, when it is used to address a node on the Internet, it will lead unambiguously to one and only one node (or to no node at all, of course, if no node with that name exists). A given domain name, assigned to a domain holder, is therefore recognized uniquely all over the world.

We Have One — and Only One — Set of IP Addresses

IP addresses are also assigned uniquely and unambiguously to Internet nodes around the world. A given IP address will lead to one and only one node (or to none). Each IP address is assigned through a system of IP address registries (RFC 7020)[3] to just one party, and is via announcement[4] recognized uniquely all over the world.

Open Standards

Having one and only one set of domain names and IP addresses enables any Internet-connected entity to identify and communicate with any other entity connected to the Internet.

But communication relies not only on the ability to convey information from one party to another; the parties must also understand each other. Both the syntax (the data format) and the semantics (meaning) of the information exchanged must be understood. The information consists of *commands* (or requests for action) and *data* (to which actions are applied). Standards ensure that all the parties interpret the commands and data in the same way (Bertin, Crespi and Magedanz 2013).

---

1 The acronym "LDH" stands for "letter digit hyphen," the three types of characters that can be used to form a syntactically correct domain name label. The use of other character types has been introduced through the "Internationalized Domain Name" program, but is beyond the scope of this paper.

2 See https://tools.ietf.org/html/rfc2826.

3 See "The Internet Numbers Registry System" (August 2013), https://tools.ietf.org/html/rfc7020.

4 "Announcement" refers to the way in which the routing protocols of the Internet distribute knowledge of where each of its nodes is located, so that IP packets can be sent from one router to the next on a path that eventually ends at the correct node.

Traditional Telecommunications

Traditional telecommunication networks involved a central "master" system with which multiple "client" systems communicated directly. The master dictated how the clients communicated. If two clients wanted to communicate with each other, they did so by first connecting individually to the master. Examples of such networks include traditional telephony and banking systems.

The standard specifying how clients could communicate was therefore defined for these networks in terms of how to communicate with the central master. Often this would be expressed in the form of an API, typically a software library designed to be incorporated into each client system.

In such a master-client arrangement, the connection from any client to the central master was cheap (in both resources and cost), but the duration-based charging scheme ensured that the total transaction cost to clients was high. A simple example is that lifting a telephone receiver and receiving a dial tone was cheap, but the cost for an actual call to another telephone was high, based on payment by the minute as the call proceeded.

Quite often, service providers developed both the master system and part or all of the client systems, and how the actual communication took place was a proprietary (commercial) secret. Providers competed with each other on both price and the efficiency of the protocol used to communicate with clients. This of course included optimizing the protocol so that production and operating costs were as low as possible. In this way, even a nominally "standard" protocol endured many proprietary modifications as providers sought competitive advantage.[5]

In such an environment, a provider could attract a competitor's customers by implementing that competitor's API (perhaps by reverse engineering or licensing), enabling communication between its own clients and those of its competitor even though they did not use the same protocol for communication. In this case the provider's master acted as a proxy between its protocol and the one used by its competitor. The same effect could of course also be achieved by an agreement between the two providers to exchange traffic between their master systems.

Internet Communication

In the Internet, end nodes can communicate directly with each other without the intervention of a central master — the network intelligence is implemented in the nodes themselves. This requires the end nodes to implement the same protocol, and for the Internet this was ensured by explicit and uniform standardization of the communication protocols in terms of "the bits on the wire." Any node that correctly implements the standards can communicate with any other node that does the same.

The difference between specifying a protocol for node-to-node communication and one for node-to-master-to-node communication may seem small — the end result is the same — but if one looks at how the specifications are developed, there is a big difference. When the protocol between nodes is specified, the development of the standard is likely to take place in an open environment in which many parties are involved. When the specification of how to communicate with a central master is developed, it is often controlled or determined by whoever owns and operates the master(s), and "clients" have little influence over the standards they are then forced to adopt.

As the Internet model of end-to-end communication has displaced the centrally controlled master-slave configurations of traditional telecommunications, the process of developing and using a protocol standard has become more *open*, in everything from participation in the creation of the standard, to access to the standard itself, to licences needed for implementation. This is no accident — an open standards process produces results that incorporate a broad range of ideas and perspectives (not just those of a single central authority); it can be implemented and used by anyone (not just those who buy into a proprietary scheme that works only within centrally controlled boundaries); and it establishes a level playing field on which competition is not distorted by proprietary advantage. Internet standards are open not because some authority decided that they should be, but because an open process produces standards that are better — technically, economically and politically — for all of the participants (ISOC 2015a).

Open Standards Development

The Internet Engineering Task Force (IETF) is often presented as a good example of a standards body that fulfills the requirements and expectations for an open standards process.[6] In 2012, along with the Institute for Electrical and Electronics Engineers (IEEE), Internet Society (ISOC), World Wide Web Consortium (W3C), and Internet Architecture Board, the IETF endorsed the "OpenStand Principles"[7] of due process, broad consensus, transparency, balance and openness. These principles codify the six key features or "abilities" that characterize and define an open standards process:

5 "To summarize, advanced intelligent network (AIN) equipment supplied by different vendors will normally not work well together. Although global network operators cannot derive any competitive advantage from different network systems of similar price and quality, they can derive such advantage from integrating these into more seamless structures and services (bundled, customized, and mass-produced) (Bohlin 1998, 109)."

6 See www.ietf.org/about/process-docs.html.

7 See https://open-stand.org/about-us/principles/.

Ability to Participate in Development of the Standard

The ability to take part in the development of a standard has two aspects: whether only certain parties can participate, and if the cost of participation is high. In many traditional telecommunication standards development processes the cost of participation is high and there is no ability to participate if you are the wrong kind of entity (regardless of how much you pay).

Ability to Access Working Documents

Even if direct participation is not possible, a standards development process might arrange for "outsiders" to review preliminary documents — perhaps because the direct participants in the process want input from others. Non-members might be interested in early drafts of a standard so that they can make earlier decisions on whether to implement the standard and assess how much the standard may affect their business. Some standards organizations do give access to all documents while others do not. For example, in the Telecommunication Standardization Sector of the International Telecommunications Union (ITU-T), some members (including Sweden) have argued that all documents should be freely available; however, a majority of ITU-T members object to free access.

Ability to Participate in Decision Making

Even where participation in a standards process is allowed (including access to working documents), it is sometimes hierarchical in that only certain membership types can participate in formal decisions (such as whether to approve a particular draft as a standard), which often are made by vote. This gives some members greater power than others to influence the final outcome of the development of a standard.

Ability to Appeal

If there is an error in the specification of a standard, or if there is a view that the specification does not solve the problem it was supposed to solve, it is essential that the process provide for appealing the decision to approve it. An appeal can lead to a change in the developed standard or to initiation of development of a new standard that replaces the old one. It can of course also be rejected.

Ability to Access the Standard

After a standard has been approved, it must be accessible to those outside of the standards development process so that it can be implemented. The business model of some standards bodies relies on control over how the standard is distributed and how much access to the standard should cost (perhaps graded according to the type of access or the type of entity seeking access). The product of an open standards process, however, must be freely available to all.

Ability to Implement the Standard

Even where access to a standard is freely available, some or all of the standard might be encumbered by intellectual property rights (such as patents) that must be licensed from their owner(s). In some cases, the licensing fee is defined by the standards organization (often an industry consortium); in other cases, it might be defined by the owner. For open standards, it is customary for the rights holder to grant an implied licence to anyone who wants to implement the standard.

STANDARDS EVOLUTION

Two examples of Internet applications for which open standards have been developed — electronic mail (email) and social media instant messaging, or "chat" — illustrate how standards evolve.

Electronic Mail

Electronic mail, with its familiar "user@example.com" addresses, is probably the most widely used and recognizable Internet service. In Internet terms, its protocols are ancient — the first email standard[8] was published in 1980. It has been updated regularly since then, maintaining backward compatibility with previous versions at each step.

Email Standards

The basic standard for email consists of two core specifications: one that specifies how an email message is formatted (with To, From and Subject headers, for example), and one that specifies how to send an email message to a recipient. The first standard for the Simple Mail Transfer Protocol (SMTP) was published in 1982.[9] It specifies a protocol in which a client opens a Transmission Control Protocol (TCP) connection to a server, interacts with the server via commands and responses and then closes the connection. Later, as email evolution led to configurations in which clients were not directly connected to servers, two more protocols — the Internet Message Access Protocol (IMAP)[10] and the Post Office Protocol (POP)[11] were developed to facilitate email retrieval.

But after evolving from direct interaction with the message store to using POP and IMAP, email clients have evolved further to become "webmail" clients, which use the World Wide Web's Hypertext Transfer Protocol (HTTP)

8 See RFC 772, https://tools.ietf.org/html/rfc772.

9 See RFC 821, https://tools.ietf.org/html/rfc821. The most recent full specification of SMTP is RFC 5321, https://tools.ietf.org/html/rfc5321.

10 See RFC 3501, https://tools.ietf.org/html/rfc3501.

11 See RFC 1939, https://tools.ietf.org/html/rfc1939.

to manage email interactions. One such client consists of a web browser that accesses a normal web page that is dynamically created by some software on the server side of the HTTP connection. This software, not run by the client, often in turn acts as an IMAP client.

Because of this, email has evolved back to a system in which the email user is connected to some application using a mechanism that is very similar to the old direct message store connections, although the connections are now made over the Web using a web client.

Standard Email Extensions

As email use expanded, the standards evolved to accommodate new demands to support non-ASCII text, images and other data formats, either as "attachments" to or in the main body of email messages. Beginning in 1991, the IETF developed a highly flexible standard for Multipurpose Internet Mail Extensions (MIME)[12] that provided support for text in character sets other than ASCII; non-text attachments such as files containing audio, video, still images and applications; and messages with multiple parts (so that a single message could include multiple text and non-text elements). It also supported privately defined extensions, so that if someone wanted to send data structured in a way known only to them, they could do so by tagging the data with a private name.[13]

MIME could have led to an explosion of private email structures, but it has not. Instead, people are trying to use a small set of common formats wherever possible: one for still images, another for video, a third for signatures and so on.

The high degree of interoperability that has been achieved by the standardization of SMTP, POP and IMAP has led to a rich marketplace of server and client software, including "webmail" that behaves as if it were an IMAP client. MIME has further enabled an extensions mechanism whereby extensions can be either standardized and interoperable or non-standardized and proprietary. This has led to a situation in which an implementer can choose from a wide variety of interoperable and proprietary email configurations.

Non-standard Email Exceptions

As we might expect of such a widely used and economically significant system, email has not evolved uniformly in the direction of interoperability. Clients today are faced with choices that go beyond the standardized IMAP and POP — they may instead choose, for example, "Exchange" or "Google." Microsoft and Google are not the only players pushing email in proprietary directions, but they serve as useful examples of the way in which market forces affect the evolution of a standard.

Microsoft Exchange

The Exchange server (and the client Outlook) can use IMAP and SMTP for communication, but for full functionality they also implement a proprietary protocol between clients and servers, which includes specific authentication mechanisms. Whether the standardized protocols should be enabled or not in the Exchange server is a configuration option, and many system administrators turn these options off. That way, if someone uses an email service that in turn is implemented with the help of Exchange, they must use the Outlook Client. It is also the case that if they want to use the Outlook Client with full functionality, they must use the Exchange server.

Google Mail

Google implements their email service by exposing both a web interface and an IMAP interface to clients. However, they use IMAP in an innovative way: to categorize mail by using tags, and exposing that to the client as folders (or containers). By tagging an email message with more than one tag, it can appear in more than one container. For this user experience to be fully realized, the client must understand Google's extension to IMAP, and many IMAP clients do indeed include support for this; however, it is not a standard IMAP way of managing tags and folders.

Social Media

The term "social media" refers to a wide range of applications and services. In this section we are interested only in the instant message, or chat, feature of most social media platforms.

Internet Relay Chat

Long ago in the time frame of the Internet — in the early 1990s — a text-only instant messaging system called Internet Relay Chat (IRC) was invented in Finland. Anyone could set up an IRC server. These service providers "peered" with each other, and the addressing was based on the name of the service plus the name of whatever was to be addressed — an individual or a chat room (or "channel"). IRC was popular and is still used by some programmers. The protocol is simple, and it is very easy to create robots that respond to messages in the various channels that in many cases act as permanent chat rooms.

IRC was defined not by a standard but by its implementation in open-source software. Anyone could look at the source code and develop either server or client software. To explain how the protocol works, an

12 See RFC 2045, https://tools.ietf.org/html/rfc2045.

13 To avoid confusion among privately defined mail extensions, the IETF defined a registry and registration procedures for message header fields in RFC 3864, https://tools.ietf.org/html/rfc3864.

experimental RFC[14] was created. But IRC still evolves as a constellation of mostly open-source implementations. Informational RFCs are released now and then explaining updates to the protocol, but they are not uniformly adopted. New features arise as enhancements to an implementation that "catch on" with other software developers, some of them coordinated by more formal groups such as the IRCv3 Working Group.

Jabber

Interest in a formally standardized chat protocol developed in the late 1990s, when the IETF launched a working group to develop one. No consensus could be reached in this working group, so instead of a single standard it published several of the contenders as experimental RFCs — the idea being to allow the community to implement the protocols, and find out from experience which one should win. Which, in fact, none of these did.

Instead, an instant messaging system called Jabber was developed outside of the IETF in 1999, as an open-source software implementation.[15] The Jabber developer community specified an eXtensible Messaging and Presence Protocol (XMPP),[16] which in 2002 was moved to, and accepted by, the IETF. Jabber began very much the way IRC did, but followed a different route; today XMPP is the subject of a full set of Internet standards overseen by an IETF working group, and is the dominant interoperable chat protocol.

Proprietary Protocols

More recently, the instant messaging and related services launched by Facebook, Twitter, Skype and Google (for example) have been based on proprietary rather than standard protocols. No description of the protocol used among the various parties that communicate is provided, thus there is no ability for a third party to implement (or even interact with) the service. Access to the Facebook service, for example, is available only from Facebook.

This model differs dramatically from the provisioning of services based on standards. To get email service or instant messaging, we can choose from a multitude of providers, each of which in turn can choose from a multitude of different providers of software — or write their own. The difference between the open development and evolution of IRC and Jabber and the current growing reliance on entirely proprietary alternatives has enormous consequences for Internet users. In a world of proprietary, non-interoperable services, users are limited to choosing either Facebook (for example) or Google — they cannot choose among alternative providers of "Facebook service" or "Google service."

MARKET FORCES

As we look at how standards have evolved, we see that the developers of software and services have cooperated in producing open standards that have led to interoperability. This has created a competitive landscape in which no single player can completely dominate the market. Thousands, if not millions, of providers of web hosting have appeared, for example, and the range of available server and client software for open-standard applications is extensive.

But providers of server software have always also had an economic interest in controlling clients, and the business models of large service providers have always favoured anti-competitive domination of a market over competition (GECON 2007). Absent an alternative countervailing value proposition based on economic advantages to these and other businesses, market forces will drive the evolution of Internet protocols and services away from interoperability and toward user "lock in."

Fortunately for users, research suggests that such value propositions in favour of openness do exist (Zhu and Zhou 2011). Two modern developments — the IoT and the "cloud" — illustrate how market forces traditionally operate against open interoperability and how they can be redirected.

The IoT

As the latest hot topic in the technology industry, the IoT has produced a mountain of commentary and analysis that runs the gamut from the breathless excitement of optimists to the dark warnings of pessimists.[17] From the standpoint of the Internet architecture, the IoT is hardly a new idea — it is simply devices connected to the Internet, each with its own IP address, just like always — but from the standpoint of our assumptions about how the Internet is used, it is indeed a radical departure.

Although it has always been "things" that are actually connected physically to the Internet, most models of interaction have taken for granted that at least one of the parties to any Internet communication is a person — a "user." In the traditional Internet, communication may be user to user, or user (client) to computer (server). The IoT adds the third option of computers talking to each other without human intervention. And this third option may involve much more than talk — after all, completely

14 See RFC 1459, https://tools.ietf.org/html/rfc1459.

15 For a history of Jabber and XMPP development compiled by the XMPP Standards Foundation, see https://xmpp.org/about/history.html, including the involvement of the IETF's XMPP Working Group, see https://tools.ietf.org/wg/xmpp/charters.

16 See https://tools.ietf.org/html/rfc6120.

17 For a summary of the "promise" and "peril" scenarios see Delic (2016).

autonomous sensor networks have been gathering and storing data without human intervention for decades. What is new in the IoT is that connected devices may also autonomously analyze the information they exchange and take actions independently as a result.

The emergence of the IoT owes more to companies' marketing incentive to make devices ever more functionally "intelligent" than to any collective sense within the Internet community that things should be able to talk to one another. What standards exist, therefore, tend to be developed by individual companies or industry consortia that focus on enhancing the capability (and therefore marketability) of the "thing" without particular regard to interoperability with "things" in other industry sectors — or with the "things" of competitors.

Case Study: Lighting Control

A simple example of an arena in which open standards are missing is control of light bulbs. Superficially, it is a simple problem to control a light bulb — the traditional method uses two wires, one live and one neutral, and a switch turns power on or off on the live wire. In modern lighting systems we can often also control the intensity and colour of the light; this can be done by using an overlay protocol on the existing wires, of course, but it is even easier to do if the light bulb is connected to a network and has a unique address, at which controlling software (acting as the light switch) can communicate with it (and perhaps many other devices) using a standard protocol.

This still sounds simple, but the problem is that there is no such standard protocol for talking to light bulbs, and no system of unique addresses for them — so no light bulb and switch interoperability, and to make matters worse, the light bulbs themselves are not interchangeable. A closer look at two examples will make this problem clearer.

Philips Hue System

In the Philips Hue system[18] the light bulb communicates with a gateway (the ZigBee Bridge) using a proprietary protocol over ZigBee,[19] a low-power digital radio technology based on the IEEE 802.15.4 standard for wireless personal area networks.[20] A light- control application communicates with the same gateway using a different proprietary IP-based protocol that is not documented by Philips, although third parties have reverse-engineered the protocol and developed libraries for a variety of programming languages (including perl, php, and python) that can be used by application developers.

A light switch must understand at least one of these two proprietary protocols — the one that runs on top of ZigBee to communicate with light bulbs, or the one that uses IP to communicate with a Hue gateway. If Philips changes the protocol, the light switch has to be updated. And, of course, unless the light switch update takes place at the same time as the Philips protocol change, there will be some interval during which the switch can't control the light. Although neither Philips nor the light switch manufacturer wants this to happen, there is no well-defined change control for the Philips protocols that includes third-party suppliers of light bulbs, switches or control applications.

LifX System

The LifX[21] system uses standard WiFi, rather than ZigBee, and runs IP directly from one device to another without an intermediate gateway. In LifX configurations the devices — light bulbs and switches, and also many other devices using the "If This Then That" web service — connect to the local wireless network and get IP addresses using Dynamic Host Configuration Protocol (DHCP). The protocol used, including the encryption, is defined by the manufacturer of the light bulb and is not publicly available. Some reverse engineering has been done to provide alternatives, but most popular access to the light bulb is via the application developed by LifX itself.

IoT Standards

The pressure for manufacturers to build "silos" — vertically integrated families of devices that talk to each other, but not to devices made by other manufacturers — is evident in this case study. Lighting control is one of the simplest and most common examples of an IoT application, and because it is a consumer-oriented technology, we would expect it to be based on standards that create interoperability, at least at the level of the simple devices (bulbs and switches) that are mass-marketed to the public. But each company imagines that its proprietary approach will become widely adopted as the "de facto" standard, with respect to which it will have an obvious competitive advantage over other companies pursuing the same "maybe it will be me" strategy. Interoperability and openness are actively detrimental to such a strategy, because they dilute the advantage that a company expects to have when "everyone" starts using its version. Consumer electronics has evolved in this way for many decades; there is no reason to expect that IoT evolution will take a different course (Blind 2004).

Only by using open standards can the light bulbs and the controlling software be made interoperable, enabling

18 See http://www2.meethue.com/en-us/about-hue/what-hue-is.

19 See http://www.zigbee.org/what-is-zigbee/.

20 Although it is not a formal standards body, the ZigBee Alliance (see http://www.zigbee.org/zigbeealliance/) is the focal point for most ZigBee technology development.

21 See www.lifx.com/.

competition that could foster innovation and evolution. Today, the lack of interoperability has severely limited the growth of IP-based connected light bulbs.

The Cloud

The term "cloud computing" refers to a shared-resource model in which individual computing devices obtain application, platform and infrastructure services such as computation, storage and software via network access to a server — or, more commonly, a distributed "cloud" of servers that collectively provide those services. In the context of this paper, we are interested in a particular feature of cloud computing: the way in which it can serve as an intermediary, or proxy, to relay communication between devices that are connected to the Internet in a way that prevents them from communicating directly with each other.

Network Address Translation

The Internet's system of global addressing supports — in principle — the end-to-end connectivity of any two Internet-connected devices. In practice, however, the most common connectivity arrangement for residential and business premises has one device — in telecom terminology, the customer premises equipment (CPE) — actually connected to an Internet access provider and all other devices at those premises connected through the CPE using network address translation (NAT). The CPE typically receives one IP address via DHCP from the access provider, and shares it with all the other devices, which do not get individual IP addresses of their own. The CPE's IP address is dynamically allocated by the service provider, so it is not associated with an Internet domain name in the DNS. And the CPE also typically acts as a firewall, filtering traffic so that all communication with its attached devices must be initiated by them.

The consequence of this arrangement is that the devices connected through such a CPE cannot actually be reached directly from the Internet, and only the CPE, with its dynamically allocated IP address, can be reached from outside. All Internet communication must therefore be initiated by the devices themselves; they communicate directly with the CPE, which uses its own IP address and a local identifier to set up the path to the other Internet-connected device and manage the subsequent communication between them.

End-to-Cloud-to-End

In such a NAT configuration, the only way that information can be exchanged between two devices is if each device opens a connection through its CPE to a server that manages the flow of data between them. Two devices configured with NAT cannot communicate directly using IP.

All "end-to-end" communication is therefore actually store-and-forward, with storage in "the cloud" as an intermediary. As cloud storage initially was created to solve the ability to communicate (or lack thereof), the specification of the protocol used does not have to be published; the cloud service is created by the same party that created the device. The communication is internal to the service, and no global communication exists. No protocol standard is needed for the (non-existing) communication.

Application Programming Interfaces

In a NAT environment, user-to-user communication is mediated by a centralized service "in the cloud." The service itself defines how to interact with it by specifying an API. This specification tells devices how to use the service, which is a very different thing from a protocol standard that specifies the way in which two users may communicate end-to-end. As the licences, terms and conditions associated with these APIs are defined by the provider of the service, the end users have little choice.

This can be viewed as a classic example of a one-sided market. For example, the service provider can change the API at any time. In practice, this will always come as a surprise to its customers, whether or not its contractual agreement with the API user says that changes will be announced before being made, or that those announcements actually are made.

Data Collection

A significant market force driving the interest in service silos defined by APIs rather than end-to-end protocols is the value of what has come to be called "big data" — collections of enormous size that have only recently become susceptible to analysis (Chen et al. 2014). With the advent of tools that make feasible calculations on entire very large data sets (rather than on smaller statistical samples), being a proxy through which communication between end users takes place has become valuable. Today we see companies just collecting data, even if they do not know what calculations to make (yet); the data sets have become valuable in themselves, creating a revenue opportunity for the service provider that in some cases can compete with the sales of the service itself.

Collecting and selling the data can also allow a service provider to lower or eliminate the fees it charges to use the service. This is naturally popular with consumers, who today in many cases enjoy the use of cloud-based services for free. But the easily recognized advantages of "free" make it harder to engage the more difficult issues of data "ownership," including access, privacy and sharing data with third parties.

AN OPEN INTERNET FUTURE

This paper has presented examples of the way in which market forces can lead to fragmentation of the nominally global and open Internet into service-oriented silos. In this concluding section we argue that the silo scenario can be avoided, and that the values of an open Internet can be extended into the future by recognizing and promoting forces that counter market forces.

The Challenge

If technical constraints (such as the Internet Protocol version 4 address length limit that led to the widespread deployment of NAT) make end-to-end communication too difficult, then users will turn to proxies that involve intermediaries in the end-to-end path. User-to-user communication via proxy introduces opportunities for third-party control, access to content and metadata, and charging. If on top of this the protocol is proprietary, then all devices must communicate with the same central cloud service. The proxy provider is in full control. From a business standpoint, of course, this sort of control is extremely valuable, and many companies today are competing vigorously to become the preferred proxy for the household.

The best-case scenario in such a third-party dominated configuration would be that devices from different manufacturers are able to communicate with and via the same proxy. But even in this case, multiple proxies may provide services for the same household. The lack of standard protocols, both between devices and between devices and cloud services, leads to the implementation of services as isolated silos. Even the APIs that define the silo services will exist for only as long as the corresponding cloud services exist. In practice, this also limits the lifetime of the devices that are sold to connect to the service, as the device itself might still be functional even if the service is turned off.

Recommendations

The protocols that have been developed within the Internet architecture are deliberately peer-to-peer. Even those that specify client-server interactions, such as the email protocols POP and IMAP, specify interactions at one level without constraining the way in which other parts of the system may be defined or implemented. Silo services define only an API that governs the entire spectrum of interaction with users. The most important recommendation for avoiding a fragmented Internet future is to promote the deployment of communication systems based on standard protocols rather than service-specific APIs.

The most broadly useful and valuable protocols are those developed by open standards processes in which everyone can participate and to which everyone can contribute. Protocols that depend on privately owned intellectual property may be subject to a variety of different licensing terms, but as with the protocols themselves, the more open the licensing terms, the more beneficial the results in the market. APIs that are specified as part of a peer-oriented (as opposed to silo-oriented) system should also be developed by an open standards process.

The gold standard for an open and transparent standards process has been set by independent organizations such as the IETF, the W3C and the IEEE, but industry alliances such as the Internet Protocol for the networking of Smart Objects Alliance or the Industrial Internet Consortium can also develop open standards. Industry-sponsored standards efforts do not always welcome the participation or contribution of the users who will be affected by their outcome, but industry leader collaboration is likely to at least minimize the number of silos and increase device interoperability for the end user.

CONCLUSION

Public sector organizations should use every opportunity that arises in procurement, regulation and project funding to require the use of open standards when they are available and to promote their development when they are not. This responsibility is especially important for socially critical systems such as electronic identification and payment schemes, for which the third-party control feature of service silos is unacceptable.

The market forces that favour service-oriented vertical integration over a disintermediated open Internet create strong economic incentives for individual companies to build silos with APIs rather than interoperable devices that implement standard protocols. Countering those forces to preserve the broad economic and social benefits of an open Internet for its users will require awareness and effort on the part of users and their public sector organizations, and a willingness to take a longer view of their business interests on the part of individual companies and industry consortia.

WORKS CITED

Bertin, Emmanuel, Noel Crespi and Thomas Magedanz, eds. 2013. "Evolution of Telecommunication Services: The Convergence of Telecom and Internet: Technologies and Ecosystems." Springer-Verl Lecture Notes in Computer Science. Heidelberg, Germany: Springer.

Blind, Knut. 2004. *The Economics of Standards: Theory, Evidence, Policy.* Cheltenham, UK: Edward Elgar Publishing.

Bohlin, Par Erik and Stanford L. Levin, eds. 1998. *Telecommunications Transformation: Technology, Strategy and Policy*. Amsterdam: IOS Press.

Chen, M., S. Mao, Y. Zhang and V. C. Leung. 2014. *Big Data — Related Technologies, Challenges and Future Prospects; Chapter 5, "Big Data Analysis."* Springer Briefs in Computer Science. Heidelberg, Germany: Springer.

Delic, Kemal A. 2016. "IoT Promises, Perils and Perspectives." *ACM Ubiquity*, February, 1–5. doi: 10.1145/2822889.

GECON. 2007. "Grid Economics and Business Models: 4th International Workshop, GECON 2007 Proceedings." Rennes, France: Springer-Verlag Lecture Notes in Computer Science.

ISOC. 2015a. "Open Internet Standards: An Internet Society Public Policy Briefing." October 30. www.internetsociety.org/sites/default/files/ISOC-PolicyBrief-OpenStandards-20151030.pdf.

Xiaoguo Zhu, Kevin and Zach Zhizhong Zhou. 2011. "Lock-In Strategy in Software Competition: Open-Source Software vs. Proprietary Software." *Information Systems Research* 23(2): 536–545. doi: 10.1287/isre.1110.0358.

CIGI PUBLICATIONS
ADVANCING POLICY IDEAS AND DEBATE

The Regime Complex for Managing Global Cyber Activities
GCIG Paper Series No. 1
Joseph S. Nye, Jr.

Tipping the Scale: An Analysis of Global Swing States in the Internet Governance Debate
GCIG Paper Series No. 2
Tim Maurer and Robert Morgus

Legal Mechanisms for Governing the Transition of Key Domain Name Functions to the Global Multi-stakeholder Community
GCIG Paper Series No. 3
Aaron Shull, Paul Twomey and Christopher S. Yoo

Legal Interoperability as a Tool for Combatting Fragmentation
GCIG Paper Series No. 4
Rolf H. Weber

Innovations in Global Governance: Toward a Distributed Internet Governance Ecosystem
GCIG Paper Series No. 5
Stefaan G. Verhulst, Beth S. Noveck, Jillian Raines and Antony Declercq

The Impact of the Dark Web on Internet Governance and Cyber Security
GCIG Paper Series No. 6
Tobby Simon and Michael Chertoff

On the Nature of the Internet
GCIG Paper Series No. 7
Leslie Daigle

Understanding Digital Intelligence and the Norms That Might Govern It
GCIG Paper Series No. 8
David Omand

ICANN: Bridging the Trust Gap
GCIG Paper Series No. 9
Emily Taylor

A Primer on Globally Harmonizing Internet Jurisdiction and Regulations
GCIG Paper Series No. 10
Michael Chertoff and Paul Rosenzweig

Connected Choices: How the Internet is Challenging Sovereign Decisions
GCIG Paper Series No. 11
Melissa E. Hathaway

Solving the International Internet Policy Coordination Problem
GCIG Paper Series No. 12
Nick Ashton-Hart

Net Neutrality: Reflections on the Current Debate
GCIG Paper Series No. 13
Pablo Bello and Juan Jung

Addressing the Impact of Data Location Regulation in Financial Services
GCIG Paper Series No. 14
James M. Kaplan and Kayvaun Rowshankish

Cyber Security and Cyber Resilience in East Africa
GCIG Paper Series No. 15
Iginio Gagliardone and Nanjira Sambuli

Global Cyberspace Is Safer than You Think: Real Trends in Cybercrime
GCIG Paper Series No. 16
Eric Jardine

The Emergence of Contention in Global Internet Governance
GCIG Paper Series No. 17
Samantha Bradshaw, Laura DeNardis, Fen Osler Hampson, Eric Jardine and Mark Raymond

Landmark EU and US Net Neutrality Decisions: How Might Pending Decisions Impact Internet Fragmentation?
GCIG Paper Series No. 18
Ben Scott, Stefan Heumann and Jan-Peter Kleinhans

The Strengths and Weaknesses of the Brazilian Internet Bill of Rights: Examining a Human Rights Framework for the Internet
GCIG Paper Series No. 19
Carolina Rossini, Francisco Brito Cruz and Danilo Doneda

The Tor Dark Net
GCIG Paper Series No. 20
Gareth Owen and Nick Savage

The Dark Web Dilemma: Tor, Anonymity and Online Policing
GCIG Paper Series No. 21
Eric Jardine

One in Three: Internet Governance and Children's Rights
GCIG Paper Series No. 22
Sonia Livingstone, John Carr and Jasmina Byrne

Combatting Cyber Threats: CSIRTs and Fostering International Cooperation on Cybersecurity
GCIG Paper Series No. 23
Samantha Bradshaw

The Privatization of Human Rights: Illusions of Consent, Automation and Neutrality
GCIG Paper Series No. 24
Emily Taylor

The Digital Trade Imbalance and Its Implications for Internet Governance
GCIG Paper Series No. 25
Susan Ariel Aaronson

A Pragmatic Approach to the Right to be Forgotten
GCIG Paper Series No. 26
Wendy Hall, Kieron O'Hara and Nigel Shadbolt

Education 3.0 and Internet Governance: A New Global Alliance for Children and Young People's Sustainable Digital Development
GCIG Paper Series No. 27
Divina Frau-Meigs and Lee Hibbard

Jurisdiction on the Internet: From Legal Arms Race to Transnational Cooperation
GCIG Paper Series No. 28
Bertrand de La Chapelle and Paul Fehlinger

Patents and Internet Standards
GCIG Paper Series No. 29
Jorge L. Contreras

Tracing the Economic Impact of Regulations on the Free Flow of Data Localization
GCIG Paper Series No. 30
Matthias Bauer, Martina F. Ferracane and Erik van der Marel

Looking Back on the First Round of New gTLD Applications: Implications for the Future of Domain Name Regulation
GCIG Paper Series No. 31
Jacqueline D. Lipton

Governance of International Trade and the Internet: Existing and Evolving Regulatory Systems
GCIG Paper Series No. 32
Harsha Vardhana Singh, Ahmed Abdel-Latif and L. Lee Tuthill

Available for free download at www.cigionline.org/publications

CIGI
Centre for International Governance Innovation
www.cigionline.org

ABOUT CIGI

The Centre for International Governance Innovation is an independent, non-partisan think tank on international governance. Led by experienced practitioners and distinguished academics, CIGI supports research, forms networks, advances policy debate and generates ideas for multilateral governance improvements. Conducting an active agenda of research, events and publications, CIGI's interdisciplinary work includes collaboration with policy, business and academic communities around the world.

CIGI's current research programs focus on three themes: the global economy; global security & politics; and international law.

CIGI was founded in 2001 by Jim Balsillie, then co-CEO of Research In Motion (BlackBerry), and collaborates with and gratefully acknowledges support from a number of strategic partners, in particular the Government of Canada and the Government of Ontario.

Le CIGI a été fondé en 2001 par Jim Balsillie, qui était alors co-chef de la direction de Research In Motion (BlackBerry). Il collabore avec de nombreux partenaires stratégiques et exprime sa reconnaissance du soutien reçu de ceux-ci, notamment de l'appui reçu du gouvernement du Canada et de celui du gouvernement de l'Ontario.

For more information, please visit www.cigionline.org.

ABOUT CHATHAM HOUSE

Chatham House, the Royal Institute of International Affairs, is based in London. Chatham House's mission is to be a world-leading source of independent analysis, informed debate and influential ideas on how to build a prosperous and secure world for all. The institute: engages governments, the private sector, civil society and its members in open debates and confidential discussions about significant developments in international affairs; produces independent and rigorous analysis of critical global, regional and country-specific challenges and opportunities; and offers new ideas to decision-makers and -shapers on how these could best be tackled from the near- to the long-term. For more information, please visit: www.chathamhouse.org.

CIGI MASTHEAD

Executive

President	Rohinton P. Medhora
Director of the International Law Research Program	Oonagh Fitzgerald
Director of the Global Security & Politics Program	Fen Osler Hampson
Director of Human Resources	Susan Hirst
Director of the Global Economy Program	Domenico Lombardi
Chief of Staff and General Counsel	Aaron Shull
Director of Communications and Digital Media	Spencer Tripp

Publications

Managing Editor, Publications	Carol Bonnett
Senior Publications Editor	Jennifer Goyder
Publications Editor	Patricia Holmes
Publications Editor	Nicole Langlois
Publications Editor	Kristen Scott Ndiaye
Publications Editor	Lynn Schellenberg
Graphic Designer	Sara Moore
Graphic Designer	Melodie Wakefield

Communications

For media enquiries, please contact communications@cigionline.org.

CIGI
67 Erb Street West
Waterloo, Ontario N2L 6C2
tel +1 519 885 2444 fax +1 519 885 5450
www.cigionline.org

CHATHAM HOUSE
The Royal Institute of International Affairs
10 St James's Square
London, England SW1Y 4LE, United Kingdom
tel +44 (0)20 7957 5700 fax +44 (0)20 7957 5710
www.chathamhouse.org